UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/071,934	02/07/2002	Markus Walther	9432-000151	9849

27572	7590	09/09/2005

HARNESS, DICKEY & PIERCE, P.L.C.
P.O. BOX 828
BLOOMFIELD HILLS, MI 48303

EXAMINER
YOUNG, DONALD G

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 09/09/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO 90C (Rev 10/03)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>03</u> MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>February 07, 2002</u>.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) <u>1-19</u> is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) <u>1-19</u> is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
    Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Objections

1.   The numbering of claims is not in accordance with 37 CFR 1.126

which requires the original numbering of the claims to be preserved

throughout the prosecution.   When claims are canceled, the remaining

claims must not be renumbered.   When new claims are presented, they

must be numbered consecutively beginning with the number next

following the highest numbered claims previously presented (whether

entered or not).

    Misnumbered claims: **(10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and**

20) have been renumbered as follows: **(9, 10, 11, 12, 13, 14, 15, 16,**

17, 18 and 19).

Claim Rejections - 35 USC § 102

2.   The following is a quotation of the appropriate paragraphs of 35

U.S.C. 102 that form the basis for the rejections under this section

made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or
> a foreign country or in public use or on sale in this country, more than one
> year prior to the date of application for patent in the United States.

3.   **Claims 1-6 and 11-16** are rejected under 35 U.S.C. 102(b) as being

anticipated by Hsu et al. (EP 1072986 A2).

    Regarding claim 1, Hsu et al. disclose a context-aware tokenizer

comprising:

- at least one context automaton module that generates a context record (contextual rules) associated with tokens of an input data stream (text sequence)(Fig. 15(a), 15(c) and paragraph [77 through 82]);

- a tokenizing automaton module having a token automaton (information extractor) that partitions (divide) said input data stream (text sequence) into predefined tokens based on pattern information contained in said token automaton and simultaneously verifying (comparing) contextual appropriateness based on said context record (paragraph [42 through 45]).

Regarding claim 2, Hsu et al. disclose a context-aware tokenizer wherein:

- said context automaton module comprises a left context automaton that populates (generate) said context record based on identified patterns that precede a given token and a right context automaton that populates (generate) said context record (contextual rules) based on identified patterns that follow said given token (Fig. 15(a), Fig. 15(c) and paragraph [77 through 82]).

Regarding claim 3, Hsu et al. disclose a context-aware tokenizer wherein:

- tokenizing automaton module maintains a data store of predefined token classes (token type)(Fig. 4 and paragraph [53]);

- assigns each token identified to at least one of said predefined token classes (paragraph [49 through 51]).

Regarding claim 4, Hsu et al. disclose a context-aware tokenizer wherein:

- tokenizer reports information indicative of the position and class membership of tokens identified (The reference teaches that Fig. 5 is the text sequence segmented into tokens using the token types listed in Fig. 4.)(Fig. 5 and paragraph [54 through 55]).

Regarding claim 5, Hsu et al. disclose a context-aware tokenizer wherein:

- tokenizing automaton defines a failure state (incorrect matches), and wherein said tokenizing automaton module monitors the occurrence of said failure state to maintain a record of the longest match (longest match corresponds to pattern results for the largest number value for $(p-n)/p+n$)) found involving said failure state to detect a default token (broader token class) in the absence of any matching patterns taken from said context automaton module (Fig. 17(a), Fig. 18 element 1810 and paragraph [83 through 87]).

Regarding claim 6, Hsu et al. disclose a context-aware tokenizer wherein:

- context automaton scans (reading) said input data stream (text sequence) in a left-to-right direction to acquire left context information and in a right-to-left direction to acquire right context information (paragraph [44 through 46]).

Regarding claim 11 (formerly claim 12), claim 11 recites the same or similar limitation as claim 1 above, and so is rejected for the same reasons.

Regarding claim 12 (formerly claim 13), claim 12 recites the same or similar limitation as claim 2 above, and so is rejected for the same reasons.

Regarding claim 13 (formerly claim 14), claim 13 recites the same or similar limitation as claim 3 above, and so is rejected for the same reasons.

Regarding claim 14 (formerly claim 15), claim 14 recites the same or similar limitation as claim 4 above, and so is rejected for the same reasons.

Regarding claim 15 (formerly claim 16), claim 15 recites the same or similar limitation as claim 5 above, and so is rejected for the same reasons.

Regarding claim 16 (formerly claim 17), claim 16 recites the same or similar limitation as claim 6 above, and so is rejected for the same reasons.

Claim Rejections - 35 USC § 103

4.   The following is a quotation of 35 U.S.C. 103(a) which forms the

basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically
> disclosed or described as set forth in section 102 of this title, if the
> differences between the subject matter sought to be patented and the prior
> art are such that the subject matter as a whole would have been obvious at
> the time the invention was made to a person having ordinary skill in the
> art to which said subject matter pertains. Patentability shall not be
> negatived by the manner in which the invention was made.

5.   **Claims 7 and 17 (formerly claim 18)** are rejected under 35

U.S.C. 103(a) as being unpatentable over Hsu et al. as applied to

claims 1 above, and in view of Reps (ACM 1998).

Regarding claims 7 and 17 (formerly claim 18), Hsu et al. fail to

teach of a tokenizer wherein said context automaton and tokenizing

automaton collectively obey a linear time operating constraint.

However, Reps does teach of a context automaton and tokenizing

automaton that collectively obeying a linear time operating constraint

(page 263 and 267). Therefore, it would have been obvious for one of

ordinary skill in the art at the time of applicant's invention to

supplement Hsu et al.'s tokenizer with Reps linear time operating

constraint to allow for reduction of storage utilization, as taught by

Reps (page 267).

6.   **Claims 8 and 18 (formerly claim 19)** are rejected under 35

U.S.C. 103(a) as being unpatentable over Hsu et al. as applied to

claim 1 above, and in view of Pereira et al. (USPN 5,781,884).

Regarding claim 8 and 18, Hsu et al. teach of an input data

stream characterized as a text string partitioned to include token

class membership information. Hsu et al. lack disclosing a text-to-speech wherein the information from the partition influences the pronunciation of the text string. However, Pereira et al. does teach of a text-to-speech synthesizer (TTS system) wherein information from said partitioned text string influences the pronunciation of said text string (col. 4, line 10 through col. 5, line 4 and col. 6, lines 20-35). Therefore, it would have been obvious to one of ordinary skill in the art at the time of applicant's invention to supplement Hsu et al.'s tokenizer with Pereira et al. text-to-speech synthesizer to allow for a multilingual system that is capable of handling a wide range of languages including Chinese or Japanese, as taught by Pereira et al. (col., lines 20-23).

7.   **Claims 9 (formerly claim 10), 10 (formerly claim 11) and 19 (formerly claim 20)** are rejected under 35 U.S.C. 103(a) as being unpatentable over Hsu et al. as applied to claim 1 above, in view of Corston-Oliver et al. (US 20020138248)

   Regarding claim 9 (formerly claim 10) and 10 (formerly claim 11), Hsu et al. fail to teach of a text processor coupled to a tokenizing automaton. However, Corston-Oliver et al. does teach of a tokenizing automaton (message parser)(Fig. 2, element 204) coupled to said text processor (linguistic analyzer) (Fig. 2, element 206) wherein input data stream (message) comprises text that lacks word unit separation symbols (Japanese) (It is well known that Japanese text does not contain word space indicators as is found in European or Romance languages). Corston-Oliver et al. also teaches said text processor

operating upon said text to identify and label multi-word

phrases/units for single unit treatment (Fig. 4, element 224 and

paragraph [50 and 88]). Therefore, it would have been obvious to one

of ordinary skill in the art at the time of applicant's invention to

supplement Hsu et al.'s tokenizer with Corston-Oliver et al.'s text

processor to allow for text to be compressed and more easily displayed

on small screens in a linguistically intelligent manner, as taught by

Corston-Oliver et al. (paragraph [1]).

Regarding claim 19 (formerly claim 20), Hsu et al. fail to teach

of generating tokenization information about input stream (message)

that includes class membership (meaning, part-of-speech) of predefined

tokens (pronoun, verb etc.) and supplying tokenization information to

a text processor. However, Corston-Oliver et al. does teach of

generating tokenization information about input stream that includes

class membership of predefined tokens and supplying tokenization

information to a text processor (linguistic analyzer) (Fig. 2, element

206) (Fig. 4, element 222, element 224 and paragraphs [25-27 and 35-

45]). Therefore, it would have been obvious to one of ordinary skill

in the art at the time of applicant's invention to supplement Hsu et

al.'s method for tokenizing with Corston-Oliver et al.'s method for

supplying tokenization information to a text processor to allow for

text to be compressed and more easily displayed on small screens in a

linguistically intelligent manner, as taught by Corston-Oliver et al.

(paragraph [1]).

Conclusion

8.   The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Kaplan (USPN 5,721,939) teaches a method and apparatus for tokenizing natural language text that minimizes required data storage and produces guaranteed incremental output. The tokenizer is in the form of a finite state transducer.

Hutchins (USPN 5,384,893) teaches a system for synthesizing a speech signal from strings of words, including a memory in which predetermined syntax tags are stored in association with entered words. A parser accesses the memory and groups the syntax tags of the entered words into phrases according to a first set of predetermined grammatical rules.

Luther (USPN 5,555,343) teaches a text parser for a text-to-speech processor that accepts a text stream and parses the text stream to detect non-spoken characters and spoken characters. A text generator generates pre-designated text sequences in response to non-spoken characters, such as special character sequences or character sequences which match format templates.

Carus (USPN 5,890,103) teaches of a tokenizing apparatus and method that includes a parser that extracts characters from the stream of text, an identifying element for identifying a token formed of characters in the stream of text that include lexical matter.

Ushioda (USPN 6,178,396) teaches of a method of attaching a token to a word class sequence whose probability of appearance in text data is equal to or more than a predetermined value.

Karaali et al. (USPN 6,182,028) teach a method, device and system to provide part-of-speech disambiguation for words. The method disambiguates the part-of-speech tags of text tokens by obtaining a set of probabilistically annotated tags for each text token, and choosing between the locally predicted tag and the alternative tag when the locally predicted tag and the alternative tag are different.

Friedman (USPN 6,182,029) teaches a computerized method for extracting information from natural-language text. The method includes parsing the text data to determine the grammatical structure of the text data and regularizing the parsed text data to form structured word terms.

Johnson et al. (USPN 6,618,722) teach a method and apparatus to make keyword selection and/or weighting as a function of a session history of user input in order to answer queries submitted by the user to a computer system by providing answers based on stored documents. The aim is to find the best answers by matching stored natural language documents both o the most recent query and to the latest query in a context that captures the recent history interaction.

Arnold et al. (USPN 6,745,161) teach a method for linguistic pattern recognition of information. Textual information is segmented into a plurality of phrases, which are then scanned for patterns of interest.

9.   Any inquiry concerning this communication or earlier communications from the examiner should be directed to Donald Young

whose telephone number is (571) 272-8134. The examiner can normally

be reached on 8:30 a.m. to 5:00 p.m..

If attempts to reach the examiner by telephone are unsuccessful,

the examiner's supervisor, Talivaldis Smits can be reached on (571)

272-7628. The fax phone number for the organization where this

application or proceeding is assigned is 571-273-8300.

10. Information regarding the status of an application may be

obtained from the Patent Application Information Retrieval (PAIR)

system. Status information for published applications may be obtained

from either Private PAIR or Public PAIR. Status information for

unpublished applications is available through Private PAIR only. For

more information about the PAIR system, see http://pair-

direct.uspto.gov. Should you have questions on access to the Private

PAIR system, contact the Electronic Business Center (EBC) at 866-217-

9197 (toll-free).

                                        Donald Young
                                        Examiner
                                        Art Unit 2654
07/28/05

                                        RICHEMOND DORVIL
                                        SUPERVISORY PATENT EXAMINER